XI

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/074,451	02/11/2002	Hyung-Bae Jeon	2013p018	8944

8791 7590 04/11/2005

BLAKELY SOKOLOFF TAYLOR & ZAFMAN
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1030

EXAMINER	
SKED, MATTHEW J	
ART UNIT	**PAPER NUMBER**
2655	

DATE MAILED: 04/11/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	10/074,451	JEON ET AL.	
	Examiner	Art Unit	
	Matthew J Sked	2655	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-12* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-12* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on _____ is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date *4/3/02*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Specification

1. The disclosure is objected to because of the following informalities: on page 3, lines 26-27, the applicant states "a method for searching signals before mixing and mixing matrix only on a condition that mixed signals are collected from a mike". It is unclear what this statement means.

Appropriate correction is required.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

3. Claims 1, 2, 10 and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kadambe (U.S. Pat. Pub. 2003/0061035A1) in view of Lee et al. (U.S. Pat. 6,424,960).

As per claims 1, 10 and 12, Kadambe teaches a blind source separation method, apparatus and computer-readable medium in a noise environment, the method comprising the steps of:

adapting the basis functions of noise signals to the present environment by using the characteristic of noise signals, which are input into a mike (optimizes an initial estimate of the mixing matrix to obtain a final mixing matrix by clustering the current

mixing samples hence adapting the initial estimate to the present environment, paragraphs 105-108);

extracting determination information for detection speech activation from the mixtures of speech signals and the mixtures of noise signals (iteratively adjusts the clustering of the mixed signal samples based upon the mixing matrix to determine the estimate for the source signals, paragraphs 109-110); and

determining a speech starting point and a speech ending point of mike signals, which come into a speech recognition unit, from the determination information (the separated source signals would inherently contain the information to detect the start and the end of speech signals, paragraphs 109-110).

Kadambe does not specifically teach or suggest using this method in voice activity detection.

However, the Examiner takes Official Notice that voice activity detection is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe for use voice activity detector because this would allow the system to turn off the speech recognizer during periods of silence hence saving power.

Kadambe does not teach training basis functions of mixed speech signals and noise signals according to a predetermined learning rule.

Lee teaches a system for classifying sources in blind signal separation that adapts the mixing matrix for each class hence training the basis functions, which are the column vectors of the mixing matrix (Fig. 5).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe to train the basis functions as taught by Lee because this would give a more robust estimate of the mixing matrix hence giving better voice detection.

4. As per claim 2, Kadambe does not teach the predetermined learning rule is independent component analysis.

Lee teaches independent component analysis is a commonly used technique for blind source separation (col. 1, lines 15-41).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe to use ICA as the predetermined learning rule because, as taught by Lee, it is a useful tool for finding structure in data and has been applied successfully to separating mixed speech signals.

5. Claims 3-9 and 11 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kadambe in view of Lee and taken in further view of Liu (U.S. Pat. 6,615,170).

As per claim 3, Kadambe teaches estimating speech and noise generation coefficients from the mixtures of noise signals and the mixtures of speech signals (iteratively adjusts the clustering of the mixed signal samples based upon the mixing matrix to determine the estimate for the source signals, paragraphs 109-110);

Kadambe and Lee do not teach computing values of likelihood of speech signals and noise signals from the speech and noise generation coefficients and computing

speech activation-determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals.

Liu teaches a method for voice detection that computes the likelihood of speech signals and noise signals from the speech and noise mixtures (col. 4, lines 32-40 and Fig. 2, elements 64, 66 and 68) and computes the speech activation determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals (calculates the ratio test statistic from the probabilities and uses this statistic in a decision function to determine if speech is present, equations 4, 7 and 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe and Lee to determine if speech is present by using the log-likelihood ratio test because, as taught by Liu, it takes into account the similarity scores of both speech and silence simultaneously hence it is more robust to background noise environments (col. 10, lines 57-63).

6. As per claim 4, Kadambe does not teach wherein the likelihood of speech signals is computed using Equation:

$$\log p\langle x \mid \theta\rangle = \log p(s) - \log(\det|A_s|)$$

where x is a mike signal, θ is a parameter, s is speech, and A_s is a mixing matrix having speech basis function information.

Lee teaches that calculating the likelihood of a data vector, given a parameter and class, is given by the equation $p(x \mid \theta, C_k) = \frac{p(s_k)}{\det[A_k]}$ (this is an equivalent equation that is not in the logarithmic form, col. 9, lines 51-57).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe to use the given equation to calculate the likelihood of speech signals as taught by Lee because it gives a simple calculation to determine the likelihood of a type of signal with the given information.

7. As per claim 5, Kadambe and Lee do not teach wherein the determination information for detecting a speech starting point is a value in which a difference between the log-likelihood of speech signals and the log-likelihood of noise signals is normalized with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal.

Liu teaches computing the speech activation determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals (calculates the ratio test statistic from the probabilities and uses this statistic in a decision function to determine if speech is present, equations 4, 7 and 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe and Lee to compute the speech activation determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals because, as taught by Liu, it takes into account the similarity scores of both speech and silence simultaneously hence it is more robust to background noise environments (col. 10, lines 57-63).

Kadambe, Lee and Liu do not teach normalizing the information for detection with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal.

However, the Examiner takes Official Notice that normalization with respect to the non-speech signal is common in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe, Lee and Liu to normalize the information for detection with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal because this would ensure that all detection information is conforming to the same noisy likelihood ratio reference.

8. As per claim 6, Kadambe and Liu do not teach a value in which a difference between the log-likelihood of speech signals and the log-likelihood of noise signals is normalized with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal, and the log-likelihood of noise signals is used as the determination information for detecting a speech starting point.

Lee teaches a value which is a difference between the log-likelihood of speech signals and the log-likelihood of noise signals and the log-likelihood of noise signals is used as the determination information for detecting a speech starting point (calculates the log-likelihood ratio text statistic that is used for speech detection and includes the log-likelihood of noise signals, equations 4, 7 and 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe and Lee to compute value which is a difference between the log-likelihood of speech signals and the log-likelihood of noise signals and the log-likelihood of noise signals is used as the determination information for detecting a speech starting point because, as taught by Liu, it takes into account the similarity scores of both speech and silence simultaneously hence it is more robust to background noise environments (col. 10, lines 57-63).

Kadambe, Lee and Liu do not teach normalizing the information for detection with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal.

However, the Examiner takes Official Notice that normalization with respect to the non-speech signal is common in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe, Lee and Liu to normalize the information for detection with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal because this would ensure that all detection information is conforming to the same noisy likelihood ratio reference.

9. As per claim 7, Kadambe and Lee do not teach wherein the determination information for detecting a speech starting point is a value in which the width of variation in a difference between the log-likelihood of speech signals and the log-likelihood of noise signals is normalized with respect to the difference between the log-likelihood of

speech signals and the log-likelihood of noise signals at the initial non-activated speech signal.

Liu teaches computing the speech activation determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals (calculates the ratio test statistic from the probabilities and uses this statistic in a decision function to determine if speech is present, equations 4, 7 and 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe and Lee to compute the speech activation determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals because, as taught by Liu, it takes into account the similarity scores of both speech and silence simultaneously hence it is more robust to background noise environments (col. 10, lines 57-63).

Kadambe, Lee and Liu do not teach the determination information is the width of variation in a difference between the likelihood of speech signals and the value of the likelihood of noise signals.

However, the Examiner takes Official Notice that the amount of variation between subsequent calculations is a notoriously well-known measure in speech detection. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe, Lee and Liu so that the determination information is the width of variation in a difference between the likelihood of speech signals and the value of the likelihood of noise signals because it would take into account spurious noise spike data hence giving better speech detection.

Kadambe, Lee and Liu do not teach normalizing the width of variation with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal.

However, the Examiner takes Official Notice that normalization with respect to the non-speech signal is common in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe, Lee and Liu to normalize the information for detection with respect to the difference between the log-likelihood of speech signals and the log-likelihood of noise signals at the initial non-activated speech signal because this would ensure that all detection information is conforming to the same noisy likelihood ratio reference.

10. As per claim 8, Kadambe and Lee do not teach the mike signals are input into a speech recognition unit in an initial mute state having noise, the state is moved into a starting point standby state when a speech starting point-determining information is greater than a first threshold value, the state is moved into a speech activation state when the speech starting point-determining information is greater than the first threshold value for a predetermined duration, the state is returned to the initial mute state when the speech starting point-determining information is not greater than the first threshold value for a predetermined duration, the state is moved into a speech ending point standby state when a speech ending point-determining information is smaller than a second threshold value in the speech activation state, the state is moved into the initial mute state when the state stays in the speech ending point standby state for more than a predetermined duration, and the state is returned to the speech activation the speech

ending point-determining information is not smaller than the second threshold value for a predetermined duration, in the step of detecting a speech starting point and a speech ending point.

Liu teaches using thresholds to determine when the system is in a speech or noise state (equation 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe and Lee to use thresholds to determine if the system is in a speech or noise state as taught by Liu because it would be a simple method to determine when speech is present hence saving processing time.

Kadambe, Lee and Liu do not teach holding the system in a standby state in between the mute and speech states until the determination information is greater than the first threshold value or less than the second threshold value for a specified duration.

However, the Examiner takes Official Notice that requiring a value to exceed a threshold for a duration to determine if the value has indeed exceeded the threshold is notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe, Lee and Liu to hold the system in a standby state in between the mute and speech states until the determination information is greater than the first threshold value or less than the second threshold value for a specified duration because it would prevent spurious noise spike data from causing speech recognition errors hence improving voice detection.

11. As per claim 9, Kadambe, Lee and Liu do not teach the first and second threshold values are determined according to the circumstance of the present noise.

However, the Examiner takes Official Notice that noise-level adaptable thresholds are notoriously well known in the art. Therefore, it would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe, Lee and Liu to have the first and second threshold values determined according to the circumstance of the present noise because it would take the current noise conditions into account in determining speech or noise hence giving better speech detection results.

12. As per claim 11, Kadambe teaches estimating speech and noise generation coefficients from the mixtures of noise signals and the mixtures of speech signals (iteratively adjusts the clustering of the mixed signal samples based upon the mixing matrix to determine the estimate for the source signals, paragraphs 109-110);

Kadambe and Lee do not teach computing values of likelihood of speech signals and noise signals from the speech and noise generation coefficients and computing speech activation-determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals.

Liu teaches a method for voice detection that computes the likelihood of speech signals and noise signals from the speech and noise mixtures (col. 4, lines 32-40 and Fig. 2, elements 64, 66 and 68) and computes the speech activation determining information from a difference between the likelihood of speech signals and the value of the likelihood of noise signals (calculates the ratio test statistic from the probabilities and

uses this statistic in a decision function to determine if speech is present, equations 4, 7 and 9).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the system of Kadambe and Lee to determine if speech is present by using the log-likelihood ratio test because, as taught by Liu, it takes into account the similarity scores of both speech and silence simultaneously hence it is more robust to background noise environments (col. 10, lines 57-63).

Conclusion

13. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Gelin (U.S. Pat. 6,327,564) and Breton (U.S. Pat. Pub. 2002/0035471A1) teach a method of speech detection that adapts the noise model. Huang et al. (U.S. Pat. Pub. 2002/0029144) and Liu et al. ("Speaker Verification Using Normalized Log-Likelihood Score") teach methods of speech detection using log-likelihood measures. Hansen et al. ("Blind Detection of Independent Dynamic Components"), Cichocki ("Blind Separation and Filtering Using State Space Models) and Bofill ("Blind Separation of More Sources Than Mixtures Using Sparsity of Their Short-Time Fourier Transform") teach methods for blind source separation.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Matthew J Sked whose telephone number is (571) 272-7627. The examiner can normally be reached on Mon-Fri (8:00 am - 4:30 pm).

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (571) 272-7628. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

MS
04/07/05

TĀLIVALDIS IVARS ŠMITS
PRIMARY EXAMINER